Interactions Between the Generalized Hadamard Product and the Eigenvalues of Symmetric Matrices

Hristo S. Sendov*

February 1, 2008

Abstract

In [8] a notion of generalized Hadamard product was introduced. We show that when certain kinds of tensors interact with the eigenvalues of symmetric matrices the resulting formulae can be nicely expressed using the generalized Hadamard product and two simple linear operations on the tensors. The Calculus-type rules developed here will be used in [9] to routinize, to a large extend, the differentiation of spectral functions. We include all the necessary definitions and results from [8] on the generalized Hadamard product to make the reading self-contained.

Keywords: spectral functions, eigenvalues, Hadamard product, tensor analysis, eigenvalue optimization, symmetric function, perturbation theory.
Mathematics Subject Classification (2000): primary: 49R50; 47A75, secondary: 15A18; 15A69.

1 Introduction

The aim of this paper is to develop some analytic tools that, we believe, will fully describe the formula for the higher order derivatives of *spectral functions* in terms of the underlying *symmetric function*. We say that a real-valued function F, on a symmetric matrix argument, is spectral if it has the following invariance property:

$$F(UXU^T) = F(X),$$

for every symmetric matrix X in its domain and every orthogonal matrix U. When U varies freely over the orthogonal matrices the invariants of the product UXU^T are the eigenvalues of the matrix

*Department of Mathematics and Statistics, University of Guelph, Guelph, Ontario, Canada N1G 2W1. Email: hssendov@uoguelph.ca. Research supported by NSERC.

X. Therefore, de facto, the function F depends only on the set of eigenvalues of X. The restriction of F to the subspace of diagonal matrices defines (almost) a function $f(x) := F(\mathrm{Diag}\, x)$ on a vector argument $x \in \mathbb{R}^n$. It is easy to see that $f : \mathbb{R}^n \to \mathbb{R}$ has the property

$$f(x) = f(Px) \text{ for any permutation matrix } P \text{ and any } x \in \operatorname{domain} f.$$

We call such functions *symmetric*. It is not difficult to see that $F(X) = f(\lambda(X))$, where $\lambda(X)$ is the vector of eigenvalues of X.

One of the main questions in the theory of spectral functions is what smoothness properties of the symmetric function f are inherited by F. The difficulties arise from the fact that the eigenvalue map, $\lambda(X)$, don't depend smoothly on its argument X. Even in domains where they are smooth, it is difficult to organize the differentiation process so that the result is in as closed form as possible.

One of the fist results in this direction (see [5]) showed that F is (continuously) differentiable at a matrix A if and only if f is, at the vector $\lambda(A)$. The formula for the gradient is compact and easy to understand. Next, [7], showed that F is twice (continuously) differentiable if, and only if, f is at $\lambda(A)$. The consideration of variety of different subcases made the differentiation laborious and the formula for the Hessian of F takes some effort to get comfortable with. Following the developments in [7], one can see that an attempt to compute the third, or higher, derivatives of F, will result in a number of subcases that will quickly become unmanageable. That is why, deriving a formula for the higher derivatives (and in the process proving that they exist) requires a language that handles all the cases in a structured way, and allows easy to work with calculus rules. In [8] we proposed such a language based on the idea of generalizing the Hadamard product between two matrices to a tensor-valued product between k matrices, $k \geq 1$. This paper is a continuation of the work there. While, [8] emphasizes on multi-linear algebra and combinatorial aspects of the generalized Hadamard product, our current work deals mainly with the calculus of the generalized Hadamard product that is related to the eigenvalues of symmetric matrices.

It is likely that high-powered analytical methods will be able to show directly that F is k times (continuously) differentiable if and only if f is, see [1]. Our approach aims to give a constructive procedure how to, knowing the k-th derivative of F, practically compute the $(k+1)$-st. The precise description of the directional expansion of the eigenvalues of a symmetric matrix, when the entries of the matrix depend only on one scalar parameter (see [3], [4]) finds far reaching practical applications in areas ranging from optimization to quantum mechanics. Formulae for the derivatives of spectral functions will naturally include, as a special case, the directional derivatives, when the symmetric matrix depends on one argument.

The paper is organized as follows. In the next section we introduce the necessary notation and definitions. The background definitions and results from [8] that will be needed are given in Section 3. All this aims to make the reading as self-contained as possible. In Section 4 we reexamine Lemma 2.4 from [7]. We distill the essential eight parts of the statement of the lemma down to two equations: the first is nothing more than a spectral decomposition of a block-diagonal symmetric matrix, while the second is a (strong) first-order expansion for dot products between (parts of) eigenvectors. The section after that contains the calculus results. The main theorems of that section are Theorem 5.1 and Theorem 5.4. They investigate how the interaction between

block-constant tensors (see below) and eigenvalues affect the generalized Hadamard product. We use these tools in [9] to derive a computable formula for the higher-order derivatives of any spectral function at a symmetric matrix X with distinct eigenvalues, as well as for the derivatives of separable spectral functions at an arbitrary symmetric matrix. (Separable spectral functions are those arising from symmetric functions $f(x) = g(x_1) + \cdots + g(x_n)$ for some function g on a scalar argument.)

2 Notation and definitions

In what follows, M^n will denote the Euclidean space of all $n \times n$ real matrices with inner product $\langle X, Y \rangle = \mathrm{tr}\,(XY^T)$, and the subspace of $n \times n$ symmetric matrices will be denoted by S^n. For $A \in S^n$, $\lambda(A) = (\lambda_1(A), ..., \lambda_n(A))$ will be the vector of its eigenvalues ordered in nonincreasing order. By O^n and P^n we will denote the set of all $n \times n$ orthogonal and permutation matrices respectively. By $\mathbb{N}_k$ we will denote the set $\{1, 2, ..., k\}$. For any vector x in $\mathbb{R}^n$, $\mathrm{Diag}\, x$ will denote the diagonal matrix with the vector x on the main diagonal, and $\mathrm{diag}\colon M^n \to \mathbb{R}^n$ will denote its conjugate operator, defined by $\mathrm{diag}\,(X) = (x_{11}, ..., x_{nn})$. By $\mathbb{R}^n_{\downarrow}$ we denote the cone of all vectors x in $\mathbb{R}^n$ such that $x_1 \geq x_2 \geq \cdots \geq x_n$. In the whole paper $\{M_m\}_{m=1}^{\infty}$ will denote a sequence of symmetric matrices converging to 0, $\{U_m\}_{m=1}^{\infty}$ will denote a sequence of orthogonal matrices. We describe sets in $\mathbb{R}^n$ and functions on $\mathbb{R}^n$ as *symmetric* if they are invariant under coordinate permutations. We denote the gradient of f by ∇f, and the Hessian by $\nabla^2 f$. In general, if the function f is k times differentiable, then $\nabla^k f(\mu)$ will denote its k-th order differential at the point μ. It can be viewed as a k-dimensional tensor on $\mathbb{R}^n$.

Definition 2.1 A *k-tensor, T, on* $\mathbb{R}^n$ is a map from $\mathbb{R}^n \times \cdots \times \mathbb{R}^n$ (k-times) to $\mathbb{R}$ that is linear in each argument separately. The set of all k-tensors on $\mathbb{R}^n$ will be denoted by $T^{k,n}$. The value of the k-tensor, T, at $(h_1, ..., h_k)$ will be denoted by $T[h_1, ..., h_k]$. The tensor is called *symmetric* if for any permutation, σ, on $\mathbb{N}_k$ it satisfies

$$T[h_{\sigma(1)}, ..., h_{\sigma(k)}] = T[h_1, ..., h_k],$$

for any $h_1, ..., h_k \in \mathbb{R}^n$.

We denote the standard basis in $\mathbb{R}^n$ by $e^1, e^2, ..., e^n$. For an arbitrary k-tensor, T, and any k-tuple of integers from $\mathbb{N}_n$, $(i_1, ..., i_k)$, we denote its $(i_1, ..., i_k)$-th element by $T^{i_1 \cdots i_k}$. (Matrices will be viewed as 2-tensors and vectors as 1-tensors.) If $T \in T^{k,n}$ and $h \in \mathbb{R}^n$, then for brevity throughout the paper, we denote by $T[h]$ the $(k-1)$-tensor on $\mathbb{R}^n$ given by $T[\cdot, ..., \cdot, h]$. Similarly for $T[M]$, when T is a k-tensor on M^n and $M \in M^n$.

For a permutation matrix $P \in P^n$ we say that $\sigma : \mathbb{N}_n \to \mathbb{N}_n$ is its corresponding permutation map and write $P \leftrightarrow \sigma$ if for any $h \in \mathbb{R}^n$ we have $Ph = (h_{\sigma(1)}, ..., h_{\sigma(n)})^T$ or, in other words, $P^T e^i = e^{\sigma(i)}$ for all $i = 1, ..., n$. The symbol δ_{ij} will denote the Kroneker delta. It is equal to one if $i = j$ and zero otherwise.

Any vector $\mu \in \mathbb{R}^n$ defines a partition of $\mathbb{N}_n$ into disjoint *blocks*, where integers i and j are in the same block if, and only if, $\mu_i = \mu_j$. Whenever μ is a vector in $\mathbb{R}^n_{\downarrow}$ we make the convention that

$$\mu_1 = \cdots = \mu_{k_1} > \mu_{k_1+1} = \cdots = \mu_{k_2} > \mu_{k_2+1} \cdots \mu_{k_r}, \qquad (k_0 = 0,\ k_r = n),$$

with corresponding partition

$$(1) \qquad I_1 := \{1, 2, ..., \iota_1\},\ I_2 := \{\iota_1 + 1, \iota_1 + 2, ..., \iota_2\}, ...,\ I_r := \{\iota_{r-1} + 1, ..., \iota_r\}.$$

For arbitrary vector μ the blocks it determines need not contain consecutive integers. Thus, we agree that the block containing the integer 1 will be the first block, I_1, the block containing the smallest integer that is not in I_1 will be the second block, I_2, and so on. This naturally enumerates all the blocks, and in general, ι_l will denote the largest integer in I_l for all $l = 1, ..., r$. Also, r will denote the number of blocks determined by μ. For any two integers, $i, j \in \mathbb{N}_n$ we will say that they are *equivalent (with respect to μ)* and write $i \sim j$ (or $i \sim_\mu j$) if $\mu_i = \mu_j$, that is, if they are in the same block. Two k-indexes $(i_1, ..., i_k)$ and $(j_1, ..., j_k)$ are called *equivalent* if $i_l \sim j_l$ for all $l = 1, 2, ..., n$, and we will write

$$(i_1, ..., i_k) \sim (j_1, ..., j_k).$$

Definition 2.2 Given a vector $\mu \in \mathbb{R}^n$, we say that a k-tensor, T, is *(μ-)block-constant* if $T^{i_1 \cdots i_k} = T^{j_1 \cdots j_k}$ whenever $(i_1, ..., i_k) \sim (j_1, ..., j_k)$.

A k-tensor valued map, $\mu \in \mathbb{R}^n \to \mathcal{F}(\mu) \in T^{k,n}$, is *block-constant* if $\mathcal{F}(\mu)$ is μ-block-constant for every μ.

The following elementary lemma motivates the definitions following it. It follows from applying the chain rule to the equality $f(\mu) = f(P\mu)$.

Lemma 2.3 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a symmetric function, k times differentiable at the point $\mu \in \mathbb{R}^n$, and let P be a permutation matrix such that $P\mu = \mu$. Then*

(i) $\nabla f(\mu) = P^T \nabla f(\mu)$,

(ii) $\nabla^2 f(\mu) = P^T \nabla^2 f(\mu) P$, *and in general*

(iii) $\nabla^s f(\mu)[h_1, ..., h_s] = \nabla^s f(\mu)[Ph_1, ..., Ph_s]$, *for any $h_1, ..., h_s \in \mathbb{R}^n$, and $s \in \mathbb{N}_k$.*

Definition 2.4 Given a vector $\mu \in \mathbb{R}^n$, we will say that $T \in T^{k,n}$ is *μ-symmetric* if for any permutation $P \in P^n$, such that $P\mu = \mu$, we have

$$T[Ph_1, ..., Ph_k] = T[h_1, ..., h_k], \text{ for any } h_1, ..., h_k \in \mathbb{R}^n.$$

A k-tensor valued map, $\mu \in \mathbb{R}^n \to \mathcal{F}(\mu) \in T^{k,n}$, is *$\mu$-symmetric* if for every $\mu \in \mathbb{R}^n$ and permutation matrix P we have

$$\mathcal{F}(P\mu)[Ph_1, ..., Ph_k] = \mathcal{F}(\mu)[h_1, ..., h_k], \text{ for any } h_1, ..., h_k \in \mathbb{R}^n.$$

Clearly, every μ-block-constant tensor is μ-symmetric, the opposite is not true. There is a slight abuse of terminology since μ-symmetric means different things for a k-tensor and for a k-tensor valued map. If the map $\mu \in \mathbb{R}^n \to \mathcal{F}(\mu) \in T^{k,n}$ is μ-symmetric, then for a fixed μ the tensor $\mathcal{F}(\mu)$ is μ-symmetric. This makes sure there will be no confusion.

By Lemma 2.3, for any differentiable enough, symmetric function $f : \mathbb{R}^n \to \mathbb{R}$ the mapping $\mu \in \mathbb{R}^n \to \nabla f(\mu) \in \mathbb{R}^n$ is a μ-symmetric, μ-block-constant, 1-tensor valued mapping. In general, for every $s \in \mathbb{N}_k$ the mapping $\mu \in \mathbb{R}^n \to \nabla^s f(\mu)$ is a μ-symmetric, s-tensor-valued map, and if continuous, then every tensor is also symmetric.

We conclude this section with the following easy lemma.

Lemma 2.5 *If a k-tensor valued map, $\mu \in \mathbb{R}^n \to T(\mu) \in T^{k,n}$, is μ-symmetric and differentiable, then its differential is also μ-symmetric.*

Proof. We use the first-order Taylor expansion formula. Let v_m be a sequence of vectors approaching zero such that $v_m/\|v_m\|$ approaches h as $m \to \infty$.

$$T(\mu + v_m)[h_1, ..., h_k] = T(\mu)[h_1, ..., h_k] + \nabla T(\mu)[h_1, ..., h_k, v_m] + o(\|v_m\|).$$

On the other hand, for any permutation P we have

$$\begin{aligned} T(\mu + v_m)[h_1, ..., h_k] &= T(P\mu + Pv_m)[Ph_1, ..., Ph_k] \\ &= T(P\mu)[Ph_1, ..., Ph_k] + \nabla T(P\mu)[Ph_1, ..., Ph_k, Pv_m] + o(\|Pv_m\|) \\ &= T(\mu)[h_1, ..., h_k] + \nabla T(P\mu)[Ph_1, ..., Ph_k, Pv_m] + o(\|v_m\|). \end{aligned}$$

Subtracting the two equalities, dividing by $\|v_m\|$ and letting m go to infinity, we get

$$\nabla T(P\mu)[Ph_1, ..., Ph_k, Ph] = T(\mu)[h_1, ..., h_k, h].$$

Since the vectors h_1,...,h_k, and h are arbitrary, the result follows. ■

3 Generalized Hadamard product

In this section we will quote briefly several definitions and results from [8] that are crucial for the development in this work. Recall that the Hadamard product of two matrices $A = [A^{ij}]$ and $B = [B^{ij}]$ of the same size is the matrix of their element-wise product $A \circ B = [A^{ij}B^{ij}]$. The standard basis on the space M^n is given by the set $\{H_{pq} \in M^n \,|\, H_{pq}^{ij} = \delta_{ip}\delta_{jq} \text{ for all } i, j \in \mathbb{N}_n\}$, where δ_{ij} is the Kronecker delta function, equal to one if $i = j$, and zero otherwise.

Definition 3.1 For each permutation σ on $\mathbb{N}_k$, we define *σ-Hadamard* product between k matrices to be a k-tensor on $\mathbb{R}^n$ as follows. Given any k basic matrices $H_{p_1q_1}$, $H_{p_2q_2}$,...,$H_{p_kq_k}$

$$(H_{p_1q_1} \circ_\sigma H_{p_2q_2} \circ_\sigma \cdots \circ_\sigma H_{p_kq_k})^{i_1i_2\ldots i_k} = \begin{cases} 1, & \text{if } i_s = p_s = q_{\sigma(s)}, \forall s = 1, ..., k, \\ 0, & \text{otherwise.} \end{cases}$$

Extend this product to a multi-linear map on k matrix arguments:

$$(H_1 \circ_\sigma H_2 \circ_\sigma \cdots \circ_\sigma H_k)^{i_1 i_2 \dots i_k} = H_1^{i_1 i_{\sigma^{-1}(1)}} \cdots H_k^{i_k i_{\sigma^{-1}(k)}}. \tag{2}$$

Let T be an arbitrary k-tensor on $\mathbb{R}^n$ and let σ be a permutation on $\mathbb{N}_k$. We define $\mathrm{Diag}^\sigma T$ to be a $2k$-tensor on $\mathbb{R}^n$ in the following way

$$(\mathrm{Diag}^\sigma T)^{i_1 \dots i_k}_{j_1 \dots j_k} = \begin{cases} T^{i_1 \dots i_k}, & \text{if } i_s = j_{\sigma(s)}, \forall s = 1, ..., k, \\ 0, & \text{otherwise.} \end{cases}$$

Notice that any $2k$-tensor, T, on $\mathbb{R}^n$ can naturally be viewed as a k-tensor on M^n in the following way

$$T[H_1, ..., H_k] = \sum_{p_1, q_1 = 1}^{n} \cdots \sum_{p_k, q_k = 1}^{n} T^{p_1 \dots p_k}_{q_1 \dots q_k} H_1^{p_1 q_1} \cdots H_k^{p_k q_k}.$$

Define dot product between two tensors in $T^{k,n}$ in the usual way:

$$\langle T_1, T_2 \rangle = \sum_{p_1, \dots, p_k = 1}^{n} T_1^{p_1 \dots p_k} T_2^{p_1 \dots p_k}.$$

We define an action (called *conjugation*) of the orthogonal group O^n on the space of all k-tensors on $\mathbb{R}^n$. For any k-tensor, T, and $U \in O^n$ this action will be denoted by $UTU^T \in T^{k,n}$:

$$(UTU^T)^{i_1 \dots i_k} = \sum_{p_1 = 1}^{n} \cdots \sum_{p_k = 1}^{n} \Big(T^{p_1 \dots p_k} U^{i_1 p_1} \cdots U^{i_k p_k} \Big). \tag{3}$$

It was shown in [8] that this action is norm preserving and associative: $V(UTU^T)V^T = (VU)T(VU)^T$ for all $U, V \in O^n$.

The Diag^σ operator, the σ-Hadamard product, and conjugation by an orthogonal matrix are connected by the following formula, see [8].

Theorem 3.2 *For any k-tensor T, any matrices H_1,...,H_k, any orthogonal matrix V, and any permutation σ in P^k we have the identity*

$$\langle T, \tilde{H}_1 \circ_\sigma \cdots \circ_\sigma \tilde{H}_k \rangle = \big(V (\mathrm{Diag}^\sigma T) V^T \big)[H_1, ..., H_k], \tag{4}$$

where $\tilde{H}_i = V^T H_i V$, $i = 1, ..., k$.

Lemma 3.3 *Let T be a k-tensor on $\mathbb{R}^n$, and H be a matrix in M^n. Let $H_{i_1 j_1}$,...,$H_{i_{k-1} j_{k-1}}$ be basic matrices in M^n, and let σ be a permutation on $\mathbb{N}_k$. Then the following identities hold.*

(i) If $\sigma^{-1}(k) = k$, then

$$\langle T, H_{i_1 j_1} \circ_\sigma \cdots \circ_\sigma H_{i_{k-1} j_{k-1}} \circ_\sigma H \rangle = \Big(\prod_{t=1}^{k-1} \delta_{i_t j_{\sigma(t)}} \Big) \sum_{t=1}^{n} T^{i_1 \dots i_{k-1} t} H^{tt}.$$

(ii) If $\sigma^{-1}(k) = l$, where $l \neq k$, then

$$\langle T, H_{i_1 j_1} \circ_\sigma \cdots \circ_\sigma H_{i_{k-1} j_{k-1}} \circ_\sigma H \rangle = \Big(\prod_{\substack{t=1 \\ t \neq l}}^{k-1} \delta_{i_t j_{\sigma(t)}} \Big) T^{i_1 \dots i_{k-1} j_{\sigma(k)}} H^{j_{\sigma(k)} i_{\sigma^{-1}(k)}}.$$

4 A refinement of a perturbation result for eigenvectors

The main limiting tool in [7] was Lemma 2.4. The statement of the lemma was broken down into nine different parts, and that lead to the consideration of variety of cases when deriving the formula for the Hessian of spectral functions. For general situations, that we aim to tackle later, such case studies will quickly become unmanageable. That is why the goal of this section is to transform Lemma 2.4 from [7] into a form more suitable for computations. We begin with a lemma (the proof is a simple combination of Lemma 5.10 in [6] and Theorem 3.12 in [2]) that will alow us to define some of the necessary notation.

Lemma 4.1 *Let $\mu \in \mathbb{R}^n_\downarrow$ and let (1) be the partition defined by μ. For any sequence of symmetric matrices $M_m \to 0$ we have that*

$$\lambda(\operatorname{Diag} \mu + M_m)^T = \mu^T + \big(\lambda(X_1^T M_m X_1)^T, ..., \lambda(X_r^T M_m X_r)^T \big)^T + o(\|M_m\|), \tag{5}$$

where $X_l := [e^{k_{l-1}+1}, ..., e^{k_l}]$, for all $l = 1, ..., r$.

We need some additional notation that will be used later as well. For a fixed $\mu \in \mathbb{R}^n$ and any square matrix H, we define

$$H^{ij}_{\text{in}} = \begin{cases} H^{ij}, & \text{if } i \sim j, \\ 0, & \text{otherwise}, \end{cases}$$
$$H^{ij}_{\text{out}} = \begin{cases} H^{ij}, & \text{if } i \not\sim j, \\ 0, & \text{otherwise}. \end{cases}$$

In other words, H_{in} extracts the diagonal blocks from the matrix H and puts zeros everywhere else, while H_{out} extracts the off diagonal blocks and fills the diagonal blocks with zeros. Clearly for any matrix H we have

$$H = H_{\text{in}} + H_{\text{out}}.$$

Throughout the whole paper, we denote

$$h_m := \big(\lambda(X_1^T M_m X_1)^T, ..., \lambda(X_r^T M_m X_r)^T \big)^T. \tag{6}$$

If also $M_m / \|M_m\|$ converges to M as m goes to infinity, since the eigenvalues are continuous functions, we can define

$$h := \lim_{m \to \infty} \frac{h_m}{\|M_m\|} = \big(\lambda(X_1^T M X_1)^T, ..., \lambda(X_r^T M X_r)^T \big)^T. \tag{7}$$

We reserve the symbols h_m and h to denote the above two vectors throughout the paper, unless stated otherwise. With this notation Lemma 4.1 says that if $M_m \to 0$, then

$$\lambda(\mathrm{Diag}\,\mu + M_m)^T = \mu^T + h_m + o(\|M_m\|). \tag{8}$$

Below is the main result of this section.

Theorem 4.2 *Let $\{M_m\}$ be a sequence of symmetric matrices converging to 0, such that $M_m/\|M_m\|$ converges to M. Let μ be in $\mathbb{R}^n_{\downarrow}$ and $U_m \to U \in O^n$ be a sequence of orthogonal matrices such that*

$$\mathrm{Diag}\,\mu + M_m = U_m\big(\mathrm{Diag}\,\lambda(\mathrm{Diag}\,\mu + M_m)\big)U_m^T, \quad \textit{for all} \;\; m = 1, 2, ...$$

Then:

(i) The orthogonal matrix U has the form

$$U = \begin{pmatrix} V_1 & 0 & \cdots & 0 \\ 0 & V_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & V_r \end{pmatrix},$$

where V_l is an orthogonal matrix with dimensions $|I_l| \times |I_l|$ for all l.

(ii) The following identity holds

$$U^T M_{\mathrm{in}} U = \mathrm{Diag}\,h, \tag{9}$$

(iii) For any indexes $i \in I_l$, $j \in I_s$, and $t \in \{1, ..., r\}$ we have the (strong) first-order expansion

$$\sum_{p \in I_t} U_m^{ip} U_m^{jp} = \delta_{ij}\delta_{lt} + \frac{\delta_{lt} - \delta_{st}}{\mu_i - \mu_j} M^{ij}\|M_m\| + o(\|M_m\|), \tag{10}$$

with the understanding that the fraction is zero whenever $\delta_{lt} = \delta_{st}$ no matter what the denominator is.

Proof. This lemma is essentially Lemma 2.4 in [7]. Indeed, Part (i) is [7, Lemma 2.4 Part (i)], and Part (ii) is an aggregate version of Parts (iv) and (vii) from there as well. To prove Part (iii) we consider several cases.

Case 1. If $i = j \in I_l$ and $t = l$, then Formula (10) becomes $\sum_{p \in I_l} \big(U_m^{ip}\big)^2 = 1 + o(\|M_m\|)$, which is exactly Part (ii) of Lemma 2.4 in [7].

Case 2. If $i \neq j \in I_l$ and $t = l$, then Formula (10) becomes $\sum_{p \in I_l} \big(U_m^{ip}\big)^2 = o(\|M_m\|)$, which is exactly Part (vi) of Lemma 2.4 in [7].

Case 3. If $i \neq j \in I_l$ and $t \neq l$, then Formula (10) becomes $\sum_{p \in I_t} U_m^{ip} U_m^{jp} = o(\|M_m\|)$, which is a consequence of Part (v) of Lemma 2.4 in [7].

Case 4. If $i \in I_l$, $j \in I_s$, with $l \neq s \neq t \neq l$, then Formula (10) becomes $\sum_{p \in I_t} U_m^{ip} U_m^{jp} = o(\|M_m\|)$, which is a consequence of Part (viii) of Lemma 2.4 in [7].

Case 5. If $i \in I_l$, $j \in I_s$, with $l \neq s$ and $t = l$, then Formula (10) becomes

$$\sum_{p \in I_t} U_m^{ip} U_m^{jp} = \frac{1}{\mu_i - \mu_j} M^{ij} \|M_m\| + o(\|M_m\|).$$

This formula requires a proof. It will be presented together with the proof of the next, last, case below.

Case 6. If $i \in I_l$, $j \in I_s$, with $l \neq s$ and $t = s$, then Formula (10) becomes

$$\sum_{p \in I_t} U_m^{ip} U_m^{jp} = -\frac{1}{\mu_i - \mu_j} M^{ij} \|M_m\| + o(\|M_m\|).$$

We now show that the expressions in both Case 5 and Case 6 are valid. Recall that Part (ix) from Lemma 2.4 in [7] says that in case when $i \in I_l$, $j \in I_s$ with $l \neq s$, we have

$$\lim_{m \to \infty} \left(\mu_{k_l} \frac{\sum_{p \in I_l} U_m^{ip} U_m^{jp}}{\|M_m\|} + \mu_{k_s} \frac{\sum_{p \in I_s} U_m^{ip} U_m^{jp}}{\|M_m\|} \right) = M^{ij}.$$

Introduce the notation

$$\beta_m^l := \frac{\sum_{p \in I_l} U_m^{ip} U_m^{jp}}{\|M_m\|}, \qquad \text{for all} \quad l = 1, 2, ..., r,$$

and notice that

$$\sum_{l=1}^{r} \beta_m^l = 0, \quad \text{for all } \ m,$$

because U_m is an orthogonal matrix and the numerator of the above sum is the product of its i-th and the j-th row. Next, Case 4 above says that

$$\lim_{m \to \infty} \sum_{t \neq l, s} \beta_m^t = 0,$$

so

$$\lim_{m \to \infty} (\beta_m^l + \beta_m^s) = 0.$$

For arbitrary reals a and b we compute

$$(a\beta^l_m + b\beta^s_m) - \frac{a-b}{\mu_{k_l} - \mu_{k_s}}\left(\mu_{k_l}\beta^l_m + \mu_{k_s}\beta^s_m\right) = (\beta^l_m + \beta^s_m)\frac{b\mu_{k_l} - a\mu_{k_s}}{\mu_{k_l} - \mu_{k_s}} \to 0,$$

as $m \to \infty$. This shows that

$$\lim_{m\to\infty}(a\beta^l_m + b\beta^s_m) = \frac{a-b}{\mu_{k_l} - \mu_{k_s}} M^{ij}.$$

When $(a,b) = (1,0)$ we obtain Case 5, and when $(a,b) = (0,1)$ we obtain Case 6. ■

5 Interactions between tensors and eigenvalues

The interactions that we will investigate between the types of tensors defined in Section 2 and the eigenvalues of symmetric matrices lead naturally to two families of linear maps. Each of the next two subsections focuses on one of these families and explains how it arises.

5.1 A family of linear maps: divided differences

Fix a vector $\mu \in \mathbb{R}^n$. In what follows, the equivalence relation between numbers from $\mathbb{N}_n$ will be determined by vector μ. We define k linear maps

$$T \in T^{k,n} \to T^{(l)}_{\mathrm{out}} \in T^{k+1,n}, \text{ for } l = 1, 2, ..., k.$$

as follows:

(11) $$\left(T^{(l)}_{\mathrm{out}}\right)^{i_1\dots i_k i_{k+1}} = \begin{cases} 0, & \text{if } i_l \sim i_{k+1}, \\ \dfrac{T^{i_1\dots i_{l-1}i_{k+1}i_{l+1}\dots i_k} - T^{i_1\dots i_{l-1}i_l i_{l+1}\dots i_k}}{\mu_{i_{k+1}} - \mu_{i_l}}, & \text{if } i_l \not\sim i_{k+1}. \end{cases}$$

Notice that if T is a μ-block-constant tensor, then so is $T^{(l)}_{\mathrm{out}}$ for each $l = 1, ..., k$. The easy-to-check claim that these maps are linear means that for any two tensors $T_1, T_2 \in T^{k,n}$ and $\alpha, \beta \in \mathbb{R}$ we have

(12) $$(\alpha T_1 + \beta T_2)^{(l)}_{\mathrm{out}} = \alpha (T_1)^{(l)}_{\mathrm{out}} + \beta (T_2)^{(l)}_{\mathrm{out}}, \quad \text{for all } l = 1, ..., k.$$

One can of course iterate this definition: on the space $T^{k+1,n}$ we can define $k+1$ liner maps into $T^{k+2,n}$, and so on. We will need a way to keep track of that chain process somehow. A good enumerating tool for our needs turns out to be the set of permutations on $\mathbb{N}_k$, $\mathbb{N}_{k+1}$,....

Given a permutation σ on $\mathbb{N}_k$ we can naturally view it as a permutation on $\mathbb{N}_{k+1}$ fixing the last element. Let τ_l be the transposition $(l, k+1)$, for all $l = 1, ..., k, k+1$. Define $k+1$ permutations, $\sigma_{(l)}$, on $\mathbb{N}_{k+1}$, as follows:

(13) $$\sigma_{(l)} = \sigma\tau_l, \text{ for } l = 1, ..., k, k+1.$$

Informally speaking, given the cycle decomposition of σ, we obtain $\sigma_{(l)}$, for each $l = 1, ..., k$, by inserting the element $k+1$ immediately after the element l, and when $l = k+1$, the permutation $\sigma_{(k+1)}$ fixes the element $k+1$. Clearly $\sigma_{(l)}^{-1}(k+1) = l$ for all l, and

$$\{\text{All permutations on } \mathbb{N}_{k+1}\} = \{\sigma\tau_l \,|\, \sigma \text{ is a permutation on } \mathbb{N}_k,\ \ l = 1, ..., k, k+1\}.$$

Theorem 5.1 *Let $\{M_m\}$ be a sequence of symmetric matrices converging to 0, such that $M_m/\|M_m\|$ converges to M. Let μ be in $\mathbb{R}^n_{\downarrow}$ and $U_m \to U \in O^n$ be a sequence of orthogonal matrices such that*

$$\operatorname{Diag}\mu + M_m = U_m\big(\operatorname{Diag}\lambda(\operatorname{Diag}\mu + M_m)\big)U_m^T, \quad \textit{for all}\ \ m = 1, 2, ...$$

Then for every block-constant k-tensor T on $\mathbb{R}^n$, any matrices H_1,...,H_k, and any permutation σ on $\mathbb{N}_k$ we have

$$\lim_{m\to\infty}\Big(\frac{U_m(\operatorname{Diag}{}^{\sigma}T)U_m^T - \operatorname{Diag}{}^{\sigma}T}{\|M_m\|}\Big)[H_1, ..., H_k] = \sum_{l=1}^{k}\big(\operatorname{Diag}{}^{\sigma_{(l)}}T^{(l)}_{\text{out}}\big)[H_1, ..., H_k, M_{\text{out}}]. \tag{14}$$

Proof. Both sides of Equation (19) are linear in each argument H_s. That is why it is enough to prove the result when H_s, for $s = 1, ..., k$, is an arbitrary matrix, $H_{i_s j_s}$, from the standard basis on M^n. Since we have that

$$\Big(U_m(\operatorname{Diag}{}^{\sigma}T)U_m^T - \operatorname{Diag}{}^{\sigma}T\Big)[H_{i_1 j_1}, ..., H_{i_k j_k}] = (U_m(\operatorname{Diag}{}^{\sigma}T)U_m^T)^{i_1...i_k}_{j_1...j_k} - (\operatorname{Diag}{}^{\sigma}T)^{i_1...i_k}_{j_1...j_k},$$

we begin by developing the first term on the right-hand side.

By the definition of the conjugate action and the fact that T is block-constant, we have

$$\begin{aligned}
(U_m(\operatorname{Diag}{}^{\sigma}T)U_m^T)^{i_1...i_k}_{j_1...j_k} &= \sum_{\substack{p_\eta, q_\eta = 1\\ \eta=1,...,k}}^{n,...,n} (\operatorname{Diag}{}^{\sigma}T)^{p_1...p_k}_{q_1...q_k}\prod_{\nu=1}^{k} U_m^{i_\nu p_\nu}U_m^{j_\nu q_\nu}\\
&= \sum_{\substack{p_\eta = 1\\ \eta=1,...,k}}^{n,...,n} T^{p_1...p_k}\prod_{\nu=1}^{k} U_m^{i_\nu p_\nu}U_m^{j_\nu p_{\sigma^{-1}(\nu)}}\\
&= \sum_{\substack{p_\eta = 1\\ \eta=1,...,k}}^{n,...,n} T^{p_1...p_k}\prod_{\nu=1}^{k} U_m^{i_\nu p_\nu}U_m^{j_{\sigma(\nu)} p_\nu}\\
&= \sum_{\substack{t_\eta = 1\\ \eta=1,...,k}}^{r,...,r} T^{\iota_{t_1}...\iota_{t_k}}\prod_{\nu=1}^{k}\Big(\sum_{p_\nu\in I_{t_\nu}} U_m^{i_\nu p_\nu}U_m^{j_{\sigma(\nu)} p_\nu}\Big).
\end{aligned}$$

Thus, we have to take the limit as m approaches infinity of the expression:

$$\frac{(U_m(\operatorname{Diag}{}^{\sigma}T)U_m^T - \operatorname{Diag}{}^{\sigma}T)^{i_1...i_k}_{j_1...j_k}}{\|M_m\|} = \frac{\displaystyle\sum_{t_1,...,t_k=1}^{r,...,r} T^{\iota_{t_1}...\iota_{t_k}}\prod_{\nu=1}^{k}\Big(\sum_{p_\nu\in I_{t_\nu}} U_m^{i_\nu p_\nu}U_m^{j_{\sigma(\nu)} p_\nu}\Big) - (\operatorname{Diag}{}^{\sigma}T)^{i_1...i_k}_{j_1...j_k}}{\|M_m\|}.$$

Assume that $i_l \in I_{v_l}$ and $j_{\sigma(l)} \in I_{s_l}$ for all $l = 1, ..., k$.

We investigate several possibilities. Suppose first that among the pairs

$$(i_1, j_{\sigma(1)}), (i_2, j_{\sigma(2)}), ..., (i_k, j_{\sigma(k)}) \tag{15}$$

at least two have nonequal entries. It will become clear, that without loss of generality we may assume they are $(i_1, j_{\sigma(1)})$ and $(i_2, j_{\sigma(2)})$, that is, $i_1 \neq j_{\sigma(1)}$ and $i_2 \neq j_{\sigma(2)}$. Using Expansion (10), for any t_1, t_2 we observe that:

$$\begin{aligned}
&\lim_{m\to\infty} \frac{1}{\|M_m\|}\Big(\sum_{p_1 \in I_{t_1}} U_m^{i_1 p_1} U_m^{j_{\sigma(1)} p_1}\Big)\Big(\sum_{p_2 \in I_{t_2}} U_m^{i_2 p_2} U_m^{j_{\sigma(2)} p_2}\Big) \\
&\quad = \lim_{m\to\infty} \frac{1}{\|M_m\|}\Big(\delta_{i_1 j_{\sigma(1)}}\delta_{v_1 t_1} + \frac{\delta_{v_1 t_1} - \delta_{s_1 t_1}}{\mu_{i_1} - \mu_{j_{\sigma(1)}}} M^{i_1 j_{\sigma(1)}}\|M_m\| + o(\|M_m\|)\Big)\times \\
&\qquad\qquad\qquad \Big(\delta_{i_2 j_{\sigma(2)}}\delta_{v_2 t_2} + \frac{\delta_{v_2 t_2} - \delta_{s_2 t_2}}{\mu_{i_2} - \mu_{j_{\sigma(2)}}} M^{i_2 j_{\sigma(2)}}\|M_m\| + o(\|M_m\|)\Big) \\
&\quad = \lim_{m\to\infty} \frac{1}{\|M_m\|}\Big(\frac{\delta_{v_1 t_1} - \delta_{s_1 t_1}}{\mu_{i_1} - \mu_{j_{\sigma(1)}}} M^{i_1 j_{\sigma(1)}}\|M_m\| + o(\|M_m\|)\Big)\Big(\frac{\delta_{v_2 t_2} - \delta_{s_2 t_2}}{\mu_{i_2} - \mu_{j_{\sigma(2)}}} M^{i_2 j_{\sigma(2)}}\|M_m\| + o(\|M_m\|)\Big) \\
&\quad = 0.
\end{aligned}$$

Since in this case by definition $(\mathrm{Diag}\,^{\sigma}T)^{i_1 \dots i_k}_{j_1 \dots j_k} = 0$ we see that the whole limit above in zero.

Suppose now, that exactly one pair has unequal entries and let it be $(i_l, j_{\sigma(l)})$. We consider two subcases depending on whether or not i_l and $j_{\sigma(l)}$ are in the same block.

If both i_l and $j_{\sigma(l)}$ are in one block, that is $v_l = s_l$, then using Expansion (10), for arbitrary t, we obtain:

$$\begin{aligned}
\lim_{m\to\infty} \frac{1}{\|M_m\|}\Big(\sum_{p \in I_t} U_m^{i_l p} U_m^{j_{\sigma(l)} p}\Big) &= \lim_{m\to\infty} \frac{1}{\|M_m\|}\Big(\delta_{i_l j_{\sigma(l)}}\delta_{v_l t} + \frac{\delta_{v_l t} - \delta_{s_l t}}{\mu_{i_l} - \mu_{j_{\sigma(l)}}} M^{i_l j_{\sigma(l)}}\|M_m\| + o(\|M_m\|)\Big) \\
&= \lim_{m\to\infty} \frac{o(\|M_m\|)}{\|M_m\|} \\
&= 0.
\end{aligned}$$

In this subcase we again have $(\mathrm{Diag}\,^{\sigma}T)^{i_1 \dots i_k}_{j_1 \dots j_k} = 0$, thus the whole limit above is zero.

If i_l and $j_{\sigma(l)}$ are in different blocks, $v_l \neq s_l$, then $(\mathrm{Diag}\,^{\sigma}T)^{i_1 \dots i_k}_{j_1 \dots j_k} = 0$ and by Expansion (10) we obtain:

$$\begin{aligned}
&\lim_{m\to\infty} \frac{1}{\|M_m\|}\Big(\sum_{t_1, \dots, t_k = 1}^{r, \dots, r} T^{\iota_{t_1} \dots \iota_{t_k}} \prod_{\nu=1}^{k}\Big(\sum_{p_\nu \in I_{t_\nu}} U_m^{i_\nu p_\nu} U_m^{j_{\sigma(\nu)} p_\nu}\Big)\Big) \\
&\qquad = \lim_{m\to\infty} \frac{1}{\|M_m\|}\Big(\sum_{t_1, \dots, t_k = 1}^{r, \dots, r} T^{\iota_{t_1} \dots \iota_{t_k}} \prod_{\nu=1}^{k}\Big(\delta_{i_\nu j_{\sigma(\nu)}}\delta_{v_\nu t_\nu} + \frac{\delta_{v_\nu t_\nu} - \delta_{s_\nu t_\nu}}{\mu_{i_\nu} - \mu_{j_{\sigma(\nu)}}} M^{i_\nu j_{\sigma(\nu)}}\|M_m\| + o(\|M_m\|)\Big)\Big).
\end{aligned}$$

We show that the limit of at most two terms in the above sum are non-zero. Indeed, summands corresponding to k-tuples $(t_1, ..., t_k)$ with $t_l \notin \{v_l, s_l\}$ are equal to zero, because $\delta_{i_l j_{\sigma(l)}} = 0$, $\delta_{v_l t_l} = \delta_{s_l t_l} = 0$, and therefore

$$\delta_{i_l j_{\sigma(l)}} \delta_{v_l t_l} + \frac{\delta_{v_l t_l} - \delta_{s_l t_l}}{\mu_{i_l} - \mu_{j_{\sigma(l)}}} M^{i_l j_{\sigma(l)}} \|M_m\| + o(\|M_m\|) = o(\|M_m\|).$$

Similarly, summands corresponding to k-tuples $(t_1, ..., t_k)$ with $t_\nu \neq v_\nu$ for some $\nu \neq l$ are equal to zero, since then $\delta_{v_\nu t_\nu} = \delta_{s_\nu t_\nu} = 0$ (recall that $v_\nu = s_\nu$ for all $\nu \neq l$). Thus, the summands with possible nonzero limit correspond to the k-tuples $(v_1, ..., v_{l-1}, v_l, v_{l+1}, ..., v_k)$ and $(v_1, ..., v_{l-1}, s_l, v_{l+1}, ..., v_k)$. On the other hand, if $t_\nu = v_\nu (= s_\nu)$ for some $\nu \neq l$, then

$$\delta_{i_\nu j_{\sigma(\nu)}} \delta_{v_\nu t_\nu} + \frac{\delta_{v_\nu t_\nu} - \delta_{s_\nu t_\nu}}{\mu_{i_\nu} - \mu_{j_{\sigma(\nu)}}} M^{i_\nu j_{\sigma(\nu)}} \|M_m\| + o(\|M_m\|) = 1 + o(\|M_m\|).$$

Thus, we can calculate that the above limit is equal to

$$\begin{aligned} \frac{T^{\iota_{v_1} \dots \iota_{v_{l-1}} \iota_{v_l} \iota_{v_{l+1}} \dots \iota_{v_k}} - T^{\iota_{v_1} \dots \iota_{v_{l-1}} \iota_{s_l} \iota_{v_{l+1}} \dots \iota_{v_k}}}{\mu_{i_l} - \mu_{j_{\sigma(l)}}} M^{i_l j_{\sigma(l)}} &= \frac{T^{i_1 \dots i_{l-1} i_l i_{l+1} \dots i_k} - T^{i_1 \dots i_{l-1} j_{\sigma(l)} i_{l+1} \dots i_k}}{\mu_{i_l} - \mu_{j_{\sigma(l)}}} M^{i_l j_{\sigma(l)}} \\ &= \frac{T^{i_1 \dots i_{l-1} i_l i_{l+1} \dots i_k} - T^{i_1 \dots i_{l-1} j_{\sigma(l)} i_{l+1} \dots i_k}}{\mu_{i_l} - \mu_{j_{\sigma(l)}}} M_{\text{out}}^{i_l j_{\sigma(l)}} \end{aligned}$$

where the first equality follows from the block-constant structure of T and the second from the premise in this case that i_l and $j_{\sigma(l)}$ are in different blocks.

In the last case when $i_\nu = j_{\sigma(\nu)}$ for all $\nu = 1, ..., k$, the limit is equal to

$$\lim_{m \to \infty} \frac{1}{\|M_m\|} \Big(T^{i_1 \dots i_k} (1 + o(\|M_m\|)) - T^{i_1 \dots i_k} \Big) = 0.$$

With that we finished calculating the limit in the left-hand side of Equation (19).

We now compute the right-hand side of Equation (19) and compare with the results above. Suppose that $\sigma(l) = m$, then by the definition of $\sigma_{(l)}$ we have $\sigma_{(l)}^{-1}(m) = k+1$, $\sigma_{(l)}^{-1}(k+1) = l$, and for any integer $i \in \mathbb{N}_{k+1} \backslash \{k+1, m\}$ we have $\sigma_{(l)}^{-1}(i) = \sigma^{-1}(i)$. Below we use the standard notation that a "hat" above a term in a product means that the term is omitted. Since $\sigma_{(l)}^{-1}(k+1) = l \neq k+1$ we use the second part of Lemma 3.3 to compute:

$$\begin{aligned} \sum_{l=1}^{k} \big(\mathrm{Diag}^{\sigma_{(l)}} T_{\text{out}}^{(l)} \big) [H_{i_1 j_1}, ..., H_{i_k j_k}, M_{\text{out}}] &= \sum_{l=1}^{k} \langle T_{\text{out}}^{(l)}, H_{i_1 j_1} \circ_{\sigma_{(l)}} \cdots \circ_{\sigma_{(l)}} H_{i_k j_k} \circ_{\sigma_{(l)}} M_{\text{out}} \rangle \\ &= \sum_{l=1}^{k} \big(T_{\text{out}}^{(l)} \big)^{i_1 \dots i_k j_{\sigma_{(l)}(k+1)}} \big(\delta_{i_1 j_{\sigma_{(l)}(1)}} \cdots \widehat{\delta_{i_l j_{\sigma_{(l)}(l)}}} \cdots \delta_{i_k j_{\sigma_{(l)}(k)}} \big) M_{\text{out}}^{j_{\sigma_{(l)}(k+1)} i_{\sigma_{(l)}^{-1}(k+1)}} \\ &= \sum_{l=1}^{k} \big(T_{\text{out}}^{(l)} \big)^{i_1 \dots i_k j_{\sigma(l)}} \big(\delta_{i_1 j_{\sigma(1)}} \cdots \widehat{\delta_{i_l j_{\sigma(l)}}} \cdots \delta_{i_k j_{\sigma(k)}} \big) M_{\text{out}}^{j_{\sigma(l)} i_l}. \end{aligned}$$

It is clear that if at least two of the pairs $(i_1, j_{\sigma(1)}), (i_2, j_{\sigma(2)}), ..., (i_k, j_{\sigma(k)})$ have different entries, then the sum is zero. Let now exactly on of the pairs have unequal entries, say $i_l \neq j_{\sigma(l)}$, then the above sum will be equal to

$$\big(T^{(l)}_{\text{out}}\big)^{i_1 \dots i_k j_{\sigma(l)}} \big(\delta_{i_1 j_{\sigma(1)}} \cdots \widehat{\delta_{i_l j_{\sigma(l)}}} \cdots \delta_{i_k j_{\sigma(k)}}\big) M^{j_{\sigma(l)} i_l}_{\text{out}}.$$

If i_l and $j_{\sigma(l)}$ are in the same block, then $\big(T^{(l)}_{\text{out}}\big)^{i_1 \dots i_k j_{\sigma(l)}} = 0$ by the definition of $T^{(l)}_{\text{out}}$. If i_l and $j_{\sigma(l)}$ are not in the same block, then the last expression above is equal to

$$\big(T^{(l)}_{\text{out}}\big)^{i_1 \dots i_k j_{\sigma(l)}} M^{j_{\sigma(l)} i_l}_{\text{out}} = \frac{T^{i_1 \dots i_{l-1} i_l i_{l+1} \dots i_k} - T^{i_1 \dots i_{l-1} j_{\sigma(l)} i_{l+1} \dots i_k}}{\mu_{i_l} - \mu_{j_{\sigma(l)}}} M^{i_l j_{\sigma(l)}}_{\text{out}},$$

because M is a symmetric matrix. Finally, if $i_\nu = j_{\sigma(\nu)}$ for all $\nu = 1, ..., k$, then again $\big(T^{(l)}_{\text{out}}\big)^{i_1 \dots i_k j_{\sigma(l)}} = 0$. These outcomes are equal to the results in the corresponding cases in the first part of the proof, the theorem follows. ■

5.2 A second family of linear maps: "inflating" diagonal hyper-planes

Recall that τ_l denotes the transposition $(l, k+1)$ on $\mathbb{N}_{k+1}$. Fix a vector $\mu \in \mathbb{R}^n$ defining the equivalence relation on $\mathbb{N}_n$. We define k linear maps

$$T \in T^{k,n} \to T^{(l)}_{\text{in}} \in T^{k+1,n}, \text{ for } l = 1, 2, ..., k.$$

as follows:

$$\big(T^{\tau_l}_{\text{in}}\big)^{i_1 \dots i_k i_{k+1}} = \begin{cases} T^{i_1 \dots i_{l-1} i_{k+1} i_{l+1} \dots i_k}, & \text{if } i_l \sim i_{k+1}, \\ 0, & \text{if } i_l \not\sim i_{k+1}. \end{cases} \tag{16}$$

Notice that if T is a block-constant tensor, then so is $T^{\tau_l}_{\text{in}}$ for each $l = 1, ..., k$. In that case, we clearly have

$$\big(T^{\tau_l}_{\text{in}}\big)^{i_1 \dots i_k i_{k+1}} = \begin{cases} T^{i_1 \dots i_{l-1} i_l i_{l+1} \dots i_k}, & \text{if } i_l \sim i_{k+1}, \\ 0, & \text{if } i_l \not\sim i_{k+1}. \end{cases}$$

It is again easy to check that these maps are linear, that is, for any two tensors $T_1, T_2 \in T^{k,n}$ and $\alpha, \beta \in \mathbb{R}$ we have

$$(\alpha T_1 + \beta T_2)^{(l)}_{\text{in}} = \alpha (T_1)^{(l)}_{\text{in}} + \beta (T_2)^{(l)}_{\text{in}}, \quad \text{for all } l = 1, ..., k.$$

Define also

$$\big(T^{\tau_l}\big)^{i_1 \dots i_k i_{k+1}} = \begin{cases} T^{i_1 \dots i_{l-1} i_l i_{l+1} \dots i_k}, & \text{if } i_l = i_{k+1}, \\ 0, & \text{if } i_l \neq i_{k+1}. \end{cases} \tag{17}$$

In other words, T^{τ_l} is a $(k+1)$-tensor with entries off the hyper plane $i_l = i_{k+1}$ equal to zero. On the hyper plane $i_l = i_{k+1}$ we have placed the original tensor T.

Before we formulate the main result or this subsection we need two technical lemmas.

Proposition 5.2 *Let T be any $k+1$-tensor, x be any vector in $\mathbb{R}^n$, let V be any orthogonal matrix, and σ a permutation on $\mathbb{N}_k$. Then the following identity holds:*

$$V\big(\mathrm{Diag}^{\,\sigma}(T[x])\big)V^T = \big(V(\mathrm{Diag}^{\,\sigma_{(k+1)}}T)V^T\big)[V(\mathrm{Diag}\,x)V^T].$$

Proof. Let $H_{i_1j_1},...,H_{i_kj_k}$ be any k basic matrices. Recall that $\sigma_{(k+1)}(i) = \sigma(i)$ for all $i \in \mathbb{N}_k$ and $\sigma_{(k+1)}(k+1) = k+1$. Using Theorem 3.2 twice, we compute

$$\begin{aligned}
\big(V\big(\mathrm{Diag}^{\,\sigma}(T[x])\big)V^T\big)^{i_1...i_k}_{j_1...j_k} &= \big(V\big(\mathrm{Diag}^{\,\sigma}(T[x])\big)V^T\big)[H_{i_1j_1},...,H_{i_kj_k}] \\
&= \langle T[x], \tilde{H}_{i_1j_1} \circ_\sigma ... \circ_\sigma \tilde{H}_{i_kj_k}\rangle \\
&= \sum_{p_1,...,p_k=1}^{n,...,n} (T[x])^{p_1...p_k} \tilde{H}^{p_1p_{\sigma^{-1}(1)}}_{i_1j_1} \cdots \tilde{H}^{p_kp_{\sigma^{-1}(k)}}_{i_kj_k} \\
&= \sum_{p_1,...,p_k,p_{k+1}=1}^{n,...,n} T^{p_1...p_{k+1}} x^{p_{k+1}} \tilde{H}^{p_1p_{\sigma^{-1}(1)}}_{i_1j_1} \cdots \tilde{H}^{p_kp_{\sigma^{-1}(k)}}_{i_kj_k} \\
&= \sum_{p_1,...,p_k,p_{k+1}=1}^{n,...,n} T^{p_1...p_{k+1}} \tilde{H}^{p_1p_{\sigma^{-1}_{(k+1)}(1)}}_{i_1j_1} \cdots \tilde{H}^{p_kp_{\sigma^{-1}_{(k+1)}(k)}}_{i_kj_k} (\mathrm{Diag}\,x)^{p_{k+1}p_{\sigma^{-1}_{(k+1)}(k+1)}} \\
&= \langle T, \tilde{H}_{i_1j_1} \circ_{\sigma_{(k+1)}} ... \circ_{\sigma_{(k+1)}} \tilde{H}_{i_kj_k} \circ_{\sigma_{(k+1)}} \mathrm{Diag}\,x\rangle \\
&= \big(V(\mathrm{Diag}^{\,\sigma_{(k+1)}}T)V^T\big)[H_{i_1j_1},...,H_{i_kj_k}, V(\mathrm{Diag}\,x)V^T] \\
&= \big(\big(V(\mathrm{Diag}^{\,\sigma_{(k+1)}}T)V^T\big)[V(\mathrm{Diag}\,x)V^T]\big)^{i_1...i_k}_{j_1...j_k}.
\end{aligned}$$

Since these equalities hold for all $i_1...i_k$ and $j_1...j_k$ we are done. ∎

The next lemma says that for any block-constant tensor T, $\mathrm{Diag}^{\,\sigma}T$ is invariant under conjugations with a block-diagonal orthogonal matrix.

Lemma 5.3 *Let T be a block-constant k-tensor on $\mathbb{R}^n$, let $U \in O^n$ be a block-diagonal matrix (both with respect to one and the same partitioning of $\mathbb{N}_n$). Then for any permutation σ in $\mathbb{N}_k$ we have the identity*

$$U(\mathrm{Diag}^{\,\sigma}T)U^T = \mathrm{Diag}^{\,\sigma}T.$$

Proof. Let $\{I_1,...,I_r\}$ be the partitioning of the integers $\mathbb{N}_n$ that determines the block structure. Notice that $U^{ip}U^{jp} = 0$ whenever $i \not\sim j$ or $i \not\sim p$, and that $\sum_{p\in I_s} U^{ip}U^{jp} = \delta_{ij}$ whenever $i \in I_s$. Let $(i_1,...,i_k)$ be an arbitrary multi index and suppose that $i_s \in I_{\nu_s}$ for $s = 1,...,k$. We expand the left-hand side of the identity:

$$\big(U(\mathrm{Diag}^{\,\sigma}T)U^T\big)^{i_1...i_k}_{j_1...j_k} = \sum_{\substack{p_s,q_s=1\\ s=1,...,k}}^{n,...,n} (\mathrm{Diag}^{\,\sigma}T)^{p_1...p_k}_{q_1...q_k} U^{i_1p_1}U^{j_1q_1} \cdots U^{i_kp_k}U^{j_kq_k}$$

$$
\begin{aligned}
&= \sum_{p_1,\ldots,p_k=1}^{n,\ldots,n} T^{p_1\ldots p_k} U^{i_1 p_1} U^{j_1 p_{\sigma^{-1}(1)}} \cdots U^{i_k p_k} U^{j_k p_{\sigma^{-1}(k)}} \\
&= \sum_{p_1,\ldots,p_k=1}^{n,\ldots,n} T^{p_1\ldots p_k} U^{i_1 p_1} U^{j_{\sigma(1)} p_1} \cdots U^{i_k p_k} U^{j_{\sigma(k)} p_k} \\
&= \sum_{t_1,\ldots,t_k=1}^{r,\ldots,r} T^{\iota_{t_1}\ldots\iota_{t_k}} \sum_{\substack{p_l\in I_{t_l}\\ l=1,\ldots,k}} U^{i_1 p_1} U^{j_{\sigma(1)} p_1} \cdots U^{i_k p_k} U^{j_{\sigma(k)} p_k} \\
&= T^{\iota_{\nu_1}\ldots\iota_{\nu_k}} \sum_{\substack{p_l\in I_{\nu_l}\\ l=1,\ldots,k}} U^{i_1 p_1} U^{j_{\sigma(1)} p_1} \cdots U^{i_k p_k} U^{j_{\sigma(k)} p_k} \\
&= T^{\iota_{\nu_1}\ldots\iota_{\nu_k}} \delta_{i_1 j_{\sigma(1)}} \cdots \delta_{i_k j_{\sigma(k)}} \\
&= T^{i_1\ldots i_k} \delta_{i_1 j_{\sigma(1)}} \cdots \delta_{i_k j_{\sigma(k)}} \\
&= \big(\mathrm{Diag}^{\,\sigma} T\big)^{i_1\ldots i_k}_{j_1\ldots j_k}.
\end{aligned}
$$

The next to the last equality follows from the fact that T is block-constant. Since the multi index $(i_1, ..., i_k, j_1, ..., j_k)$ was arbitrary, the claim in the lemma follows. ∎

Theorem 5.4 *Let $U \in O^n$ be a block-diagonal orthogonal matrix. Let M be an arbitrary symmetric matrix, and let $h \in \mathbb{R}^n$ be a vector, such that*

$$U^T M_{\mathrm{in}} U = \mathrm{Diag}\, h. \tag{18}$$

Let H_1,...,H_k be arbitrary matrices, and let σ be a permutation on $\mathbb{N}_k$. Then

(i) for any block-constant $(k+1)$-tensor T on $\mathbb{R}^n$,

$$\langle T[h], \tilde{H}_1 \circ_\sigma \cdots \circ_\sigma \tilde{H}_k\rangle = \langle T, H_1 \circ_{\sigma_{(k+1)}} \cdots \circ_{\sigma_{(k+1)}} H_k \circ_{\sigma_{(k+1)}} M_{\mathrm{in}}\rangle$$

(ii) for any block-constant k-tensor T on $\mathbb{R}^n$

$$\langle T^{\tau_l}[h], \tilde{H}_1 \circ_\sigma \cdots \circ_\sigma \tilde{H}_k\rangle = \langle T^{\tau_l}_{\mathrm{in}}, H_1 \circ_{\sigma_{(l)}} \cdots \circ_{\sigma_{(l)}} H_k \circ_{\sigma_{(l)}} M_{\mathrm{in}}\rangle, \quad \textit{for all } l = 1, ..., k,$$

where the permutations $\sigma_{(l)}$, for $l = 0, 1, ..., k$ are defined by (13), $\tilde{H}_i = U^T H_i U$ for all $i = 1, ..., k$.

Proof. To see that the first identity holds we use Theorem 3.2, Proposition 5.2, Formula (18), and Lemma 5.3 in that order, as follows:

$$
\begin{aligned}
\langle T[h], \tilde{H}_1 \circ_\sigma \cdots \circ_\sigma \tilde{H}_k\rangle &= \big(U(\mathrm{Diag}^{\,\sigma} T[h])U^T\big)[H_1, ..., H_k] \\
&= \big(U(\mathrm{Diag}^{\,\sigma_{(k+1)}} T)U^T\big)[H_1, ..., H_k, U(\mathrm{Diag}\, h)U^T]
\end{aligned}
$$

$$
\begin{aligned}
&= \Big(U(\mathrm{Diag}^{\,\sigma_{(k+1)}}T)U^T\Big)[H_1, ..., H_k, M_{\mathrm{in}}] \\
&= \Big(\mathrm{Diag}^{\,\sigma_{(k+1)}}T\Big)[H_1, ..., H_k, M_{\mathrm{in}}] \\
&= \langle T, H_1 \circ_{\sigma_{(k+1)}} \cdots \circ_{\sigma_{(k+1)}} H_k \circ_{\sigma_{(k+1)}} M_{\mathrm{in}}\rangle.
\end{aligned}
$$

The last equality follows again from Theorem 3.2.

To show the second identity, since both sides are linear in H_s for every $s = 1, ..., k$, it is enough to prove it only in the case when H_s is an arbitrary basic matrix $H_{i_s j_s}$. Fix k basic matrices $H_{i_1j_1},...,H_{i_kj_k}$ and suppose that $i_s \in I_{\nu_s}$ for $s = 1, ..., k$. The left-hand side is equal to

$$
\begin{aligned}
\langle T^{\tau_l}[h], \tilde{H}_{i_1j_1} \circ_\sigma \cdots \circ_\sigma \tilde{H}_{i_kj_k}\rangle &= \Big(U(\mathrm{Diag}^{\,\sigma}T^{\tau_l}[h])U^T\Big)[H_{i_1j_1}, ..., H_{i_kj_k}] \\
&= \Big(U(\mathrm{Diag}^{\,\sigma}T^{\tau_l}[h])U^T\Big)^{i_1\ldots i_k}_{j_1\ldots j_k} \\
&= \sum_{\substack{p_1,\ldots,p_k=1 \\ q_1,\ldots,q_k=1}}^{n,\ldots,n} (\mathrm{Diag}^{\,\sigma}T^{\tau_l}[h])^{p_1\ldots p_k}_{q_1\ldots q_k} U^{i_1p_1}U^{j_1q_1}\cdots U^{i_kp_k}U^{j_kq_k} \\
&= \sum_{p_1,\ldots,p_k=1}^{n,\ldots,n} (T^{\tau_l}[h])^{p_1\ldots p_k} U^{i_1p_1}U^{j_1p_{\sigma^{-1}(1)}}\cdots U^{i_kp_k}U^{j_kp_{\sigma^{-1}(k)}} \\
&= \sum_{p_1,\ldots,p_k=1}^{n,\ldots,n} (T^{\tau_l}[h])^{p_1\ldots p_k} U^{i_1p_1}U^{j_{\sigma(1)}p_1}\cdots U^{i_kp_k}U^{j_{\sigma(k)}p_k} \\
&= \sum_{p_1,\ldots,p_k=1}^{n,\ldots,n} \sum_{p_{k+1}=1}^{n} (T^{\tau_l})^{p_1\ldots p_kp_{k+1}} h^{p_{k+1}} U^{i_1p_1}U^{j_{\sigma(1)}p_1}\cdots U^{i_kp_k}U^{j_{\sigma(k)}p_k} \\
&= \sum_{p_1,\ldots,p_k=1}^{n,\ldots,n} T^{p_1\ldots p_k} h^{p_l} U^{i_1p_1}U^{j_{\sigma(1)}p_1}\cdots U^{i_kp_k}U^{j_{\sigma(k)}p_k} \\
&= \sum_{s_1,\ldots,s_k=1}^{r,\ldots,r} T^{\iota_{s_1}\ldots\iota_{s_k}} \sum_{\substack{p_\eta\in I_{s_\eta} \\ \eta=1,\ldots,k}} h^{p_l} U^{i_1p_1}U^{j_{\sigma(1)}p_1}\cdots U^{i_kp_k}U^{j_{\sigma(k)}p_k} \\
&= T^{i_1\ldots i_k}\delta_{i_1j_{\sigma(1)}}\cdots\widehat{\delta_{i_lj_{\sigma(l)}}}\cdots\delta_{i_kj_{\sigma(k)}} \sum_{p_l\in I_{\nu_l}} h^{p_l}U^{i_lp_l}U^{j_{\sigma(l)}p_l} \\
&= T^{i_1\ldots i_k}\delta_{i_1j_{\sigma(1)}}\cdots\widehat{\delta_{i_lj_{\sigma(l)}}}\cdots\delta_{i_kj_{\sigma(k)}} M_{\mathrm{in}}^{i_lj_{\sigma(l)}}.
\end{aligned}
$$

Now we evaluate the right-hand side of the identity. We will use the second part of Lemma 3.3 since $\sigma_{(l)}^{-1}(k+1) = l \neq k+1$. Recall also that $\sigma_{(l)}(s) = \sigma(s)$ for $s \in \mathbb{N}_{k+1}\backslash\{l, k+1\}$ and $\sigma_{(l)}(k+1) = \sigma(l)$ for all $l = 1, ..., k$.

$$
\begin{aligned}
&\langle T^{\tau_l}_{\mathrm{in}}, H_{i_1j_1} \circ_{\sigma_{(l)}} \cdots \circ_{\sigma_{(l)}} H_{i_kj_k} \circ_{\sigma_{(l)}} M_{\mathrm{in}}\rangle \\
&\qquad = \Big(T^{\tau_l}_{\mathrm{in}}\Big)^{i_1\ldots i_kj_{\sigma_{(l)}(k+1)}} \delta_{i_1j_{\sigma_{(l)}(1)}}\cdots\widehat{\delta_{i_lj_{\sigma_{(l)}(l)}}}\cdots\delta_{i_kj_{\sigma_{(l)}(k)}} M_{\mathrm{in}}^{j_{\sigma_{(l)}(k+1)}i_{\sigma^{-1}_{(l)}(k+1)}} \\
&\qquad = \Big(T^{\tau_l}_{\mathrm{in}}\Big)^{i_1\ldots i_kj_{\sigma(l)}} \delta_{i_1j_{\sigma(1)}}\cdots\widehat{\delta_{i_lj_{k+1}}}\cdots\delta_{i_kj_{\sigma(k)}} M_{\mathrm{in}}^{j_{\sigma(l)}i_l}
\end{aligned}
$$

$$= T^{i_1\dots i_k}\delta_{i_1 j_{\sigma(1)}}\cdots\widehat{\delta_{i_l j_{k+1}}}\cdots\delta_{i_k j_{\sigma(k)}} M_{\text{in}}^{j_{\sigma(l)} i_l}$$
$$= T^{i_1\dots i_k}\delta_{i_1 j_{\sigma(1)}}\cdots\widehat{\delta_{i_l j_{k+1}}}\cdots\delta_{i_k j_{\sigma(k)}} M_{\text{in}}^{i_l j_{\sigma(l)}}$$
$$= T^{i_1\dots i_k}\delta_{i_1 j_{\sigma(1)}}\cdots\widehat{\delta_{i_l j_{\sigma(l)}}}\cdots\delta_{i_k j_{\sigma(k)}} M_{\text{in}}^{i_l j_{\sigma(l)}}.$$

In the third equality above we used the fact that T is block-constant, plus the fact that $M_{\text{in}}^{j_{\sigma(l)} i_l} = 0$ if $j_{\sigma(l)} \not\sim i_l$. In the fourth we used the fact that M is a symmetric matrix. The last equality holds because we changed the format of the missing multiple, while keeping the present multiples the same. ∎

Proposition 5.5 Let $U \in O(n)$ be an block-diagonal orthogonal matrix, let H be an arbitrary $n \times n$ matrix, and σ an arbitrary permutation on $\mathbb{N}_k$.

(i) If T is a $(k+1)$-tensor such that for some fixed $l \in \mathbb{N}_k$ we have $T^{p_1 \dots p_l \dots p_{k+1}} = 0$ whenever $p_l \sim p_{k+1}$, then
$$\big(U(\operatorname{Diag}^{\sigma(l)} T)U^T\big)[H_{\text{in}}] = 0.$$

(ii) If T is a $(k+1)$-tensor such that for some fixed $l \in \mathbb{N}_k$ we have $T^{p_1 \dots p_l \dots p_{k+1}} = 0$ whenever $p_l \not\sim p_{k+1}$, then
$$\big(U(\operatorname{Diag}^{\sigma(l)} T)U^T\big)[H_{\text{out}}] = 0.$$

(iii) If T is any $(k+1)$-tensor, then
$$\big(U(\operatorname{Diag}^{\sigma(k+1)} T)U^T\big)[H_{\text{out}}] = 0.$$

Proof. Fix an index l in $\mathbb{N}_k$. Let $H_{i_1 j_1},\dots,H_{i_k j_k}$ be arbitrary basic matrices, and let H be an arbitrary matrix. Using the definitions we compute.

$$\big(U(\operatorname{Diag}^{\sigma(l)} T)U^T\big)[H_{i_1 j_1},\dots,H_{i_k j_k}, H] = \sum_{i_{k+1}, j_{k+1}=1}^{n,n} \big(U(\operatorname{Diag}^{\sigma(l)} T)U^T\big)^{i_1\dots i_{k+1}}_{j_1 \dots j_{k+1}} H^{i_{k+1} j_{k+1}}$$
$$= \sum_{i_{k+1}, j_{k+1}=1}^{n,n} \sum_{\substack{p_s, q_s=1\\ s=1,\dots,k+1}}^{n,\dots,n} (\operatorname{Diag}^{\sigma(l)} T)^{p_1\dots p_{k+1}}_{q_1 \dots q_{k+1}} U^{i_1 p_1} U^{j_1 q_1} \cdots U^{i_{k+1} p_{k+1}} U^{j_{k+1} q_{k+1}} H^{i_{k+1} j_{k+1}}$$
$$= \sum_{i_{k+1}, j_{k+1}=1}^{n,n} \sum_{\substack{p_s=1\\ s=1,\dots,k+1}}^{n,\dots,n} T^{p_1\dots p_{k+1}} U^{i_1 p_1} U^{j_1 p_{\sigma_{(l)}^{-1}(1)}} \cdots U^{i_{k+1} p_{k+1}} U^{j_{k+1} p_{\sigma_{(l)}^{-1}(k+1)}} H^{i_{k+1} j_{k+1}}$$
$$= \sum_{i_{k+1}, j_{k+1}=1}^{n,n} \sum_{\substack{p_s=1\\ s=1,\dots,k+1}}^{n,\dots,n} T^{p_1\dots p_{k+1}} U^{i_1 p_1} U^{j_{\sigma_{(l)}(1)} p_1} \cdots U^{i_l p_l} U^{j_{\sigma_{(l)}(l)} p_l} \cdots U^{i_{k+1} p_{k+1}} U^{j_{\sigma_{(l)}(k+1)} p_{k+1}} H^{i_{k+1} j_{k+1}}$$

$$= \sum_{i_{k+1},j_{k+1}=1}^{n,n} \sum_{\substack{p_s=1\\ s=1,\dots,k+1}}^{n,\dots,n} T^{p_1\cdots p_{k+1}} U^{i_1p_1} U^{j_{\sigma(1)}p_1} \cdots U^{i_lp_l} U^{j_{k+1}p_l} \cdots U^{i_{k+1}p_{k+1}} U^{j_{\sigma(l)}p_{k+1}} H^{i_{k+1}j_{k+1}}.$$

Suppose now that T is a $(k+1)$-tensor with $T^{p_1\cdots p_l\cdots p_{k+1}} = 0$ whenever $p_l \sim p_{k+1}$ and that $H = H_{\text{in}}$. Then $H^{i_{k+1}j_{k+1}} \neq 0$ implies that $i_{k+1} \sim j_{k+1}$. In that case, by the fact that U is block diagonal, $U^{j_{k+1}p_l} U^{i_{k+1}p_{k+1}} \neq 0$ implies that $p_l \sim p_{k+1}$, which implies that $T^{p_1\cdots p_l\cdots p_{k+1}} = 0$. Thus every summand in the double sum above is zero.

In the second case, suppose T is a $(k+1)$-tensor with $T^{p_1\cdots p_l\cdots p_{k+1}} = 0$ whenever $p_l \not\sim p_{k+1}$ and $H = H_{\text{out}}$. Then $H^{i_{k+1}j_{k+1}} \neq 0$ implies that $i_{k+1} \not\sim j_{k+1}$. In that case, by the fact that U is block diagonal, $U^{j_{k+1}p_l} U^{i_{k+1}p_{k+1}} \neq 0$ implies that $p_l \not\sim p_{k+1}$, which implies that $T^{p_1\cdots p_l\cdots p_{k+1}} = 0$. The sum is zero.

In the third case, suppose that T is any $(k+1)$-tensor and $H = H_{\text{out}}$. A calculation almost identical to the above one (it differs only in the last step) shows that

$$\big(U(\operatorname{Diag}^{\sigma_{(k+1)}} T)U^T\big)[H_{i_1j_1}, \dots, H_{i_kj_k}, H] =$$
$$\sum_{i_{k+1},j_{k+1}=1}^{n,n} \sum_{\substack{p_s=1\\ s=1,\dots,k+1}}^{n,\dots,n} T^{p_1\cdots p_{k+1}} U^{i_1p_1} U^{j_{\sigma(1)}p_1} \cdots U^{i_kp_k} U^{j_{\sigma(k)}p_k} U^{i_{k+1}p_{k+1}} U^{j_{k+1}p_{k+1}} H^{i_{k+1}j_{k+1}}.$$

Then $H^{i_{k+1}j_{k+1}} \neq 0$ implies that $i_{k+1} \not\sim j_{k+1}$. In that case, by the fact that U is block diagonal, $U^{j_{k+1}p_{k+1}} U^{i_{k+1}p_{k+1}} = 0$. Again the sum is zero. ■

The next result is a consequence of Theorem 3.2, Theorem 5.4, and Proposition 5.5 applied with $U = I$ and using the fact that $M = M_{\text{in}} + M_{\text{out}}$.

Corollary 5.6 *Let $U \in O^n$ be a block-diagonal orthogonal matrix and let σ be a permutation on $\mathbb{N}_k$. Let M be an arbitrary symmetric matrix, and $h \in \mathbb{R}^n$ be a vector, such that $U^T M_{\text{in}} U = \operatorname{Diag} h$. Then*

(i) for any block-constant $(k+1)$-tensor T on $\mathbb{R}^n$,

$$U\big(\operatorname{Diag}^{\sigma}(T[h])\big)U^T = \big(\operatorname{Diag}^{\sigma_{(k+1)}} T\big)[M];$$

(ii) for any block-constant k-tensor T on $\mathbb{R}^n$

$$U\big(\operatorname{Diag}^{\sigma}(T^{\tau_l}[h])\big)U^T = \big(\operatorname{Diag}^{\sigma_{(l)}} T_{\text{in}}^{\tau_l}\big)[M], \quad \textit{for all } l = 1, \dots, k,$$

where the permutations $\sigma_{(l)}$, for $l \in \mathbb{N}_k$, are defined by (13).

Notice that if the vector μ, defining the equivalence relation on $\mathbb{N}_n$, has distinct coordinates, then every tensor from $T^{k,n}$ is block-constant and the block-diagonal orthogonal matrices are precisely the signed identity matrices (those with plus or minus one on the main diagonal and zeros everywhere else). In this case we also have $i \sim j$ if, and only if, $i = j$ and thus $T_{\text{in}}^{\tau_l} = T^{\tau_l}$. Moreover, since Proposition 5.5 holds for arbitrary matrices (symmetric or not), we obtain the next result, valid for an arbitrary matrix H.

Corollary 5.7 *Let σ be a permutation on $\mathbb{N}_k$ and let H be an arbitrary matrix. Then*

(i) for any $(k+1)$-tensor T on $\mathbb{R}^n$,

$$\mathrm{Diag}^{\,\sigma}(T[\mathrm{diag}\, H]) = \big(\mathrm{Diag}^{\,\sigma_{(k+1)}} T\big)[H];$$

(ii) for any k-tensor T on $\mathbb{R}^n$

$$\mathrm{Diag}^{\,\sigma}(T^{\tau_l}[\mathrm{diag}\, H]) = \big(\mathrm{Diag}^{\,\sigma_{(l)}} T^{\tau_l}\big)[H], \quad \textit{for all } l = 1, ..., k,$$

where the permutations $\sigma_{(l)}$, for $l \in \mathbb{N}_k$, are defined by (13).

Finally, combining Theorem 5.1 and Proposition 5.5 we get the next corollary.

Corollary 5.8 *Let $\{M_m\}$ be a sequence of symmetric matrices converging to 0, such that $M_m/\|M_m\|$ converges to M. Let μ be in $\mathbb{R}^n_{\downarrow}$ and $U_m \to U \in O^n$ be a sequence of orthogonal matrices such that*

$$\mathrm{Diag}\,\mu + M_m = U_m\big(\mathrm{Diag}\,\lambda(\mathrm{Diag}\,\mu + M_m)\big)U_m^T, \quad \textit{for all} \;\; m = 1, 2, ...$$

Then for every block-constant k-tensor T on $\mathbb{R}^n$, and any permutation σ on $\mathbb{N}_k$ we have

$$\lim_{m\to\infty} \frac{U_m(\mathrm{Diag}^{\,\sigma} T)U_m^T - \mathrm{Diag}^{\,\sigma} T}{\|M_m\|} = \sum_{l=1}^{k} \big(\mathrm{Diag}^{\,\sigma_{(l)}} T^{(l)}_{\mathrm{out}}\big)[M]. \tag{19}$$

6 A determinant connection

Let $T \in T^{k,n}$ be any μ-symmetric tensor on R^n. A little bit of thought shows that T can be decomposed into $T = A + B$, where $A, B \in T^{k,n}$, A is a block-constant tensor and B has the following property: for every multi index $(i_1, ..., i_k)$ with pairwise distinct entries, $B^{i_1 \dots i_k} = 0$. In this section we investigate a rather curious fact about tensors with the last property.

For any two vectors $x, y \in \mathbb{R}^k$, we say that y *refines* x (or that x *is refined by* y), and write $x \preceq y$, if $y_i = y_j$ implies $x_i = x_j$ for any $i, j \in \mathbb{N}_k$.

Lemma 6.1 *Let T_1, T_2,...,T_s be s-tensors on $\mathbb{R}^n$ with the following two properties:*

(i) for every multi index $(i_1, ..., i_s)$ we have

$$(T_1^{i_1 \dots i_s}, T_2^{i_1 \dots i_s}, ..., T_s^{i_1 \dots i_s}) \preceq (i_1, ..., i_s).$$

(ii) Whenever the multi index $(i_1, ..., i_s)$ has pairwise distinct entries, then $T_l^{i_1 \dots i_s} = 0$ for all l.

For any two multi indexes $(i_1,...,i_s)$, $(j_1,...,j_s)$ with entries from the set $\mathbb{N}_n$ we define the $s \times s$ matrix, $\Delta\binom{i_1...i_s}{j_1...j_s}$, as follows:

$$\Delta\binom{i_1...i_s}{j_1...j_s}^{pq} = \begin{cases} \delta_{i_p j_q}, & \text{if } q < s, \\ T_p^{i_1...i_s}\delta_{i_p j_q}, & \text{if } q = s, \end{cases}$$

then $\det(\Delta\binom{i_1...i_s}{j_1...j_s}) = 0$.

Proof. Fix a multi index $(i_1,...,i_s)$. If it has pairwise distinct entries, then by the second property, the last column in the matrix $\Delta\binom{i_1...i_s}{j_1...j_s}$ will be zero. If two of its entries i_{s_1} and i_{s_2} are equal, then by the first property, row s_1 will be equal to row s_2 and again the determinant will be zero. ■

Note 6.2 The second condition in the preceding lemma in trivially satisfied if $s > n$.

Proposition 6.3 *Suppose that T_1, T_2,...,T_s be s-tensors on $\mathbb{R}^n$ with the two properties in Lemma 6.1, then the following identity holds:*

$$\sum_{\sigma\in P^{s-1}}\sum_{l=1}^{s}\operatorname{sign}(\sigma_{(l)})\operatorname{Diag}^{\sigma_{(l)}}T_l = 0. \tag{20}$$

Proof. Fix an arbitrary multi index $\binom{i_1...i_s}{j_1...j_s}$. In the following calculation, the third equality uses the facts $P^s = \{\sigma_{(l)}\,|\,\sigma\in P^{s-1},\ l\in\mathbb{N}_s\}$ and $\sigma_{(l)}(l) = s$ for $l\in\mathbb{N}_s$, $\sigma\in P^{s-1}$.

$$\begin{aligned}\Big(\sum_{\sigma\in P^{s-1}}\sum_{l=1}^{s}\operatorname{sign}(\sigma_{(l)})\operatorname{Diag}^{\sigma_{(l)}}T_l\Big)^{i_1...i_s}_{j_1...j_s} &= \sum_{\sigma\in P^{s-1}}\sum_{l=1}^{s}\operatorname{sign}(\sigma_{(l)})T_l^{i_1...i_s}\delta_{i_1 j_{\sigma_{(l)}(1)}}\cdots\delta_{i_s j_{\sigma_{(l)}(s)}}\\ &= \det(\Delta\binom{i_1...i_s}{j_1...j_s})\\ &= 0,\end{aligned}$$

where $\Delta\binom{i_1...i_s}{j_1...j_s}$ is defined in Lemma 6.1 for the family of tensors $T_1, T_2, ..., T_s$. ■

Recall the definitions of the linear maps $T\in T^{k,n}\to T_{\text{out}}^{(l)}\in T^{k+1,n}$, for $l\in\mathbb{N}_k$. For convenience in this section we also define

$$T_{\text{out}}^{(k+1)} \equiv 0, \text{ for every } T\in T^{k,n}.$$

Lemma 6.4 *Suppose that $T\in T^{k,n}$ is a symmetric tensor such that $T^{i_1...i_k} = 0$ for any multi index $(i_1,...,i_k)$ with pairwise distinct entries. Then the $k+1$ $(k+1)$-tensors*

$$T_{\text{out}}^{(1)}, T_{\text{out}}^{(2)}, ..., T_{\text{out}}^{(k)}, T_{\text{out}}^{(k+1)}$$

satisfy the two conditions in Lemma 6.1 (with $s = k+1$).

Proof. To show the first property, fix a multi index $(i_1, ..., i_{k+1})$ and suppose that $i_p = i_q$ for some $1 \le p < q \le k+1$. If $q = k+1$, then $\left(T^{(p)}_{\text{out}}\right)^{i_1 \dots i_{k+1}} = \left(T^{(q)}_{\text{out}}\right)^{i_1 \dots i_{k+1}} = 0$. If $q < k+1$, then

$$
\begin{aligned}
\left(T^{(p)}_{\text{out}}\right)^{i_1 \dots i_{k+1}} &= \frac{T^{i_1 \dots i_{p-1} i_{k+1} i_{p+1} \dots i_{q-1} i_q i_{q+1} \dots i_k} - T^{i_1 \dots i_k}}{\mu_{i_{k+1}} - \mu_{i_p}} \\
&= \frac{T^{i_1 \dots i_{p-1} i_q i_{p+1} \dots i_{q-1} i_{k+1} i_{q+1} \dots i_k} - T^{i_1 \dots i_k}}{\mu_{i_{k+1}} - \mu_{i_p}} \\
&= \frac{T^{i_1 \dots i_{p-1} i_p i_{p+1} \dots i_{q-1} i_{k+1} i_{q+1} \dots i_k} - T^{i_1 \dots i_k}}{\mu_{i_{k+1}} - \mu_{i_q}} \\
&= \left(T^{(q)}_{\text{out}}\right)^{i_1 \dots i_{k+1}},
\end{aligned}
$$

where in the second equality we used the fact that T is symmetric, while in the third we used $i_p = i_q$.

The verification of the second condition in Lemma 6.1 follows immediately from the fact that T has that property. ■

Using the fact that $T^{(k+1)}_{\text{out}} \equiv 0$, we obtain the next corollary from Proposition 6.3.

Corollary 6.5 *Suppose that $T \in T^{k,n}$ is a symmetric tensor such that $T^{i_1 \dots i_k} = 0$ for any multi index $(i_1, ..., i_k)$ with pairwise distinct entries. Then we have the following identity:*

$$
\sum_{\sigma \in P^k} \sum_{l=1}^{k} \operatorname{sign}(\sigma_{(l)}) \operatorname{Diag}^{\sigma_{(l)}} T^{(l)}_{\text{out}} = 0. \tag{21}
$$

7 Lifting of a tensor determined by the cycle type of a permutation

The goal of this final section is to prove a result in the spirit of Corollary 5.6.

It is well known that every permutation ν on $\mathbb{N}_k$ has a unique decomposition into a product of disjoint cycles. Denote by s the number of disjoint cycles. These cycles partition the set of integers, $\mathbb{N}_k$, in a natural way: two integers $j, i \in N_k$ are in the same partition if $\nu^l(i) = j$ for some l. We enumerate the sets in the partition in the natural way: let $I_{\nu,1}$ be the set of the partition that contains the integer 1; let $I_{\nu,2}$ be the set that contains the smallest integer not in $I_{\nu,1}$; let $I_{\nu,3}$ be the set containing the smallest integer not in $I_{\nu,1} \cup I_{\nu,2}$, and so on.

Take a vector $x \in \mathbb{R}^k$. We will say that the *permutation ν refines vector x* (or that x is refined by ν), and write $x \preceq \nu$, if $x_l = x_{\nu(l)}$ for every $l = 1, 2, ..., k$. In order to know a vector x, refined by ν, it is enough to know the value of one coordinate with index from every cycle of ν. In other words, for a fixed x, refined by ν, the vector $(p_1, ..., p_s)$ defined by

$$p_l := x_i, \text{ for every } l \in \mathbb{N}_s \text{ and some } i \in I_{\nu,l},$$

completely *specifies x given ν*. The ordering on the cycles of ν, we agreed on above, makes sure that this is a one-to-one correspondence between $\mathbb{R}^s$ and the set $\{x \in \mathbb{R}^k \,|\, x \preceq \nu\}$.

Let now T be a tensor from $T^{s,n}$. (Notice that the dimension of the tensor is equal to the number of cycles of $\nu \in P^k$.) Clearly $s \leq k$ with equality if, and only if, ν is the identity permutation. We define T^ν to be a tensor from $T^{k,n}$ defined component wise by

$$(T^\nu)^{i_1 \dots i_k} = \begin{cases} T^{p_1 \dots p_s}, & \text{if } (i_1, \dots, i_k) \preceq \nu, \\ 0, & \text{otherwise}, \end{cases}$$

where $(p_1, \dots, p_s)$ is the vector that specifies $(i_1, \dots, i_k)$ given ν. Informally speaking, T^ν has the tensor T placed on the diagonal "subspace" defined by the coordinate equalities $\{i_l = i_{\nu(l)} \,|\, l \in \mathbb{N}_k\}$ and zeros every where else.

When τ_l is the transposition $(l, k+1)$ on $\mathbb{N}_{k+1}$, this definition coincides with the definition of T^{τ_l} given by Equation (17).

Fix a vector $\mu \in \mathbb{R}^n$. For the rest of this section, when a block-constant tensor or a block-diagonal matrix is mentioned, it will be understood that the blocks are determined by the vector μ as explained in Section 2. Let M be a symmetric matrix and $h \in R^n$ be such that

$$U^T M_{\text{in}} U = \operatorname{Diag} h, \tag{22}$$

for some block-diagonal orthogonal matrix $U \in O^n$.

Lemma 7.1 *For any block-constant tensor, T, on $\mathbb{R}^n$ we have*

$$T[h] = T[\operatorname{diag} M],$$

where vector h and matrix M are defined by Equation (22).

Proof. Suppose that T is a k-tensor. The proof is a direct calculation from the definitions:

$$\begin{aligned} (T[h])^{i_1 \dots i_{k-1}} &= \sum_{i_k=1}^{n} T^{i_1 \dots i_{k-1} i_k} h^{i_k} \\ &= \sum_{l=1}^{r} T^{i_1 \dots i_{k-1} \iota_l} \sum_{i \in I_l} h^i \\ &= \sum_{l=1}^{r} T^{i_1 \dots i_{k-1} \iota_l} \sum_{i \in I_l} M^{ii} \\ &= \sum_{i_k=1}^{n} T^{i_1 \dots i_{k-1} i_k} M^{i_k i_k} \\ &= (T[\operatorname{diag} M])^{i_1 \dots i_{k-1}}. \end{aligned}$$

■

Proposition 7.2 *Let ν be a permutation on $\mathbb{N}_{k+1}$ with s disjoint cycles such that $\nu(k+1) = k+1$. Then for any block-constant tensor T in $T^{s,n}$ we have the identity:*

$$T^{\nu}[h] = T^{\nu}[\operatorname{diag} M], \tag{23}$$

where vector h and matrix M are defined by Equation (22).

Proof. If the multi index $(i_1, ..., i_k, i_{k+1})$ is refined by ν, then $(i_1, ..., i_k, j)$ is also refined by ν for any $j \in \mathbb{N}_n$. Moreover, the vector that specifies it (given ν) will look like $(p_1, p_2, ..., p_{s-1}, i_{k+1})$ because $\nu(k+1) = k+1$ and the cycle containing the integer $k+1$ has one element. Therefore, as $(i_1, ..., i_k, i_{k+1})$ goes over all possible multi indexes of dimension $k+1$, refined by ν, $(p_1, p_2, ..., p_{s-1}, i_{k+1})$ will go over all multi indexes of dimension s. This correspondence is one-to-one. Fix an arbitrary multi index, $(i_1, ..., i_k, i_{k+1})$. If it is not refined by ν, then

$$(T^{\nu}[h])^{i_1 \dots i_k} = \sum_{i_{k+1}=1}^{n} (T^{\nu})^{i_1 \dots i_k i_{k+1}} h^{i_{k+1}} = 0,$$

and similarly the right-hand side of (23) is equal to zero. If the multi index is refined by ν, then using the previous lemma we compute:

$$\begin{aligned}
(T^{\nu}[h])^{i_1 \dots i_k} &= \sum_{i_{k+1}=1}^{n} (T^{\nu})^{i_1 \dots i_k i_{k+1}} h^{i_{k+1}} \\
&= \sum_{i_{k+1}=1}^{n} T^{p_1 \dots p_{s-1} i_{k+1}} h^{i_{k+1}} \\
&= (T[h])^{p_1 \dots p_{s-1}} \\
&= (T[\operatorname{diag} M])^{p_1 \dots p_{s-1}} \\
&= \sum_{i_{k+1}=1}^{n} T^{p_1 \dots p_{s-1} i_{k+1}} (\operatorname{diag} M)^{i_{k+1}} \\
&= (T^{\nu}[\operatorname{diag} M])^{i_1 \dots i_k}.
\end{aligned}$$

■

The following result complements Corollary 5.6. It follows by combining Proposition 7.2 and Corollary 5.7.

Corollary 7.3 *Let ν be a permutation on $\mathbb{N}_{k+1}$ with s disjoint cycles such that $\nu(k+1) = k+1$. Then for any block-constant tensor T in $T^{s,n}$ and permutation σ on $\mathbb{N}_k$ we have the identity:*

$$\operatorname{Diag}^{\sigma}\big(T^{\nu}[h]\big) = \big(\operatorname{Diag}^{\sigma(k+1)} T^{\nu}\big)[M], \tag{24}$$

where vector h and matrix M are defined by Equation (22).

References

[1] Y. GRABOVSKI, O. HIJAB, and I. RIVIN. Differentiability of funcitons of matrices. *World Wide Web*, http://front.math.ucdavis.edu, 2003.

[2] J.-B. HIRIART-URRUTY and D. YE. Sensitivity analysis of all eigenvalues of a symmetric matrix. *Numerische Mathematik*, 70:45–72, 1992.

[3] T. KATO. *Perturbation Theory for Linear Operators*. Grundlehren der Mathematischen Wissenschaften, Band 132. Springer-Verlag New York, Inc., Berlin, second edition edition, 1976.

[4] E.C. KEMBLE. *The Fundamental Principles of Quantum Mechanics*. Dover, New York, 1958.

[5] A.S. LEWIS. Derivatives of spectral functions. *Mathematics of Operations Research*, 21:576–588, 1996.

[6] A.S. LEWIS. Nonsmooth analysis of eigenvalues. *Mathematical Programming*, 84:1–24, 1999.

[7] A.S. LEWIS and H.S. SENDOV. Twice differentiable spectral functions. *SIAM Journal of Matrix Analysis and Applications*, 23(2):368–386, 2001.

[8] H.S. SENDOV. Generalized Hadamard product and the derivatives of spectral functions. *Research Report, Department of Mathematics and Statistics, University of Guelph, Ontario*, 2004. Mathematical Series 2004-308.

[9] H.S. SENDOV. On the higher-order derivatives of spectral functions: Two special cases. *Research Report, Department of Mathematics and Statistics, University of Guelph, Ontario*, 2004. Mathematical Series 2004-310.